

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C07K 14/29, C12N 15/86, A61K 31/70</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 98/16554</b> <b>(43) International Publication Date:</b> 23 April 1998 (23.04.98)
<b>(21) International Application Number:</b> PCT/US97/19044 <b>(22) International Filing Date:</b> 17 October 1997 (17.10.97) <b>(30) Priority Data:</b> 08/733,230 17 October 1996 (17.10.96) US <b>(71) Applicant:</b> UNIVERSITY OF FLORIDA [US/US]; 223 Grinter Hall, Gainesville, FL 32611 (US). <b>(72) Inventors:</b> BARBET, Anthony, F.; 8803 S.W. 138th Street, Archer, FL 32618 (US). GANTA, Roman, Reddy; 6519 West Newberry Road #802, Gainesville, FL 32605 (US). MCGUIRE, Travis, C.; S.W. 920 Crestview, Pullman, WA 99163 (US). BURRIDGE, Michael, J.; 10021 S.W. 67th Drive, Gainesville, FL 32608 (US). NYIKA, Aceme; House 6282, Unit J, Chitungwiza, Harare (ZW). RURANGIRWA, Fred, R.; 2065 N.W. Turner Drive, Pullman, WA 99163 (US). MAHAN, Suman, M.; 71 Olwyn Avenue, Strathaven, Harare (ZW). <b>(74) Agents:</b> SALIWANCHIK, David, R. et al.; Saliwanchik, Lloyd & Saliwanchik, Suite A-1, 2421 N.W. 41st Street, Gainesville, FL 32606-6669 (US).		<b>(81) Designated States:</b> AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GE, HU, ID, IL, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SD, SG, SI, SK, SL, TR, TT, UA, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> NUCLEIC ACID VACCINES AGAINST RICKETTSIAL DISEASES AND METHODS OF USE <b>(57) Abstract</b>  Described are nucleic acid vaccines containing genes to protect animals or humans against rickettsial diseases. Also described are polypeptides and methods of using these polypeptides to detect antibodies to pathogens.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## DESCRIPTION

### NUCLEIC ACID VACCINES AGAINST RICKETTSIAL DISEASES AND METHODS OF USE

5

This invention was made with government support under USAID Grant No. LAG-1328-G-00-3030-00. The government has certain rights in this invention.

#### Cross-Reference to a Related Application

10

This is a continuation-in-part of U.S. patent application Serial No. 08/733,230, filed October 17, 1996.

#### Technical Field

15

This invention relates to nucleic acid vaccines for rickettsial diseases of animals, including humans.

#### Background of the Invention

20

The rickettsias are a group of small bacteria commonly transmitted by arthropod vectors to man and animals, in which they may cause serious disease. The pathogens causing human rickettsial diseases include the agent of epidemic typhus, *Rickettsia prowazekii*, which has resulted in the deaths of millions of people during wartime and natural disasters. The causative agents of spotted fever, e.g., *Rickettsia rickettsii* and *Rickettsia conorii*, are also included within this group. Recently, new types of human rickettsial disease caused by members of the tribe *Ehrlichiae* have been described. *Ehrlichiae* infect leukocytes and endothelial cells of many different mammalian species, some of them causing serious human and veterinary diseases. Over 400 cases of human ehrlichiosis, including some fatalities, caused by *Ehrlichia chaffeensis* have now been reported. Clinical signs of human ehrlichiosis are similar to those of Rocky Mountain spotted fever, including fever, nausea, vomiting, headache, and rash.

25

30

Heartwater is another infectious disease caused by a rickettsial pathogen, namely *Cowdria ruminantium*, and is transmitted by ticks of the genus *Amblyomma*. The disease occurs throughout most of Africa and has an estimated endemic area of about 5 million square miles. In endemic areas, heartwater is a latent infection in indigenous breeds of cattle that have been subjected to centuries of natural selection. The problems occur where the disease contacts susceptible or naive cattle and other ruminants. Heartwater has been confirmed to be on the

island of Guadeloupe in the Caribbean and is spreading through the Caribbean Islands. The tick vectors responsible for spreading this disease are already present on the American mainland and threaten the livestock industry in North and South America.

In acute cases of heartwater, animals exhibit a sudden rise in temperature, signs of anorexia, cessation of rumination, and nervous symptoms including staggering, muscle twitching, and convulsions. Death usually occurs during these convulsions. Peracute cases of the disease occur where the animal collapses and dies in convulsions having shown no preliminary symptoms. Mortality is high in susceptible animals. Angora sheep infected with the disease have a 90% mortality rate while susceptible cattle strains have up to a 60% mortality rate.

If detected early, tetracycline or chloramphenicol treatment are effective against rickettsial infections, but symptoms are similar to numerous other infections and there are no satisfactory diagnostic tests (Helmick, C., K. Bernard, L. D'Angelo [1984] *J. Infect. Dis.* 150:480).

Animals which have recovered from heartwater are resistant to further homologous, and in some cases heterologous, strain challenge. It has similarly been found that persons recovering from a rickettsial infection may develop a solid and lasting immunity. Individuals recovered from natural infections are often immune to multiple isolates and even species. For example, guinea pigs immunized with a recombinant *R. conorii* protein were partially protected even against *R. rickettsii* (Vishwanath, S., G. McDonald, N. Watkins [1990] *Infect. Immun.* 58:646). It is known that there is structural variation in rickettsial antigens between different geographical isolates. Thus, a functional recombinant vaccine against multiple isolates would need to contain multiple epitopes, e.g., protective T and B cell epitopes, shared between isolates. It is believed that serum antibodies do not play a significant role in the mechanism of immunity against rickettsia (Uilenberg, G. [1983] *Advances in Vet. Sci. and Comp. Med.* 27:427-480; Du Plessis, Plessis, J.L. [1970] *Onderstepoort J. Vet. Res.* 37(3):147-150).

Vaccines based on inactivated or attenuated rickettsiae have been developed against certain rickettsial diseases, for example against *R. prowazekii* and *R. rickettsii*. However, these vaccines have major problems or disadvantages, including undesirable toxic reactions, difficulty in standardization, and expense (Woodward, T. [1981] "Rickettsial diseases: certain unsettled problems in their historical perspective," In *Rickettsia and Rickettsial Diseases*, W. Burgdorfer and R. Anacker, eds., Academic Press, New York, pp. 17-40).

A vaccine currently used in the control of heartwater is composed of live infected sheep blood. This vaccine also has several disadvantages. First, expertise is required for the

intravenous inoculation techniques required to administer this vaccine. Second, vaccinated animals may experience shock and so require daily monitoring for a period after vaccination. There is a possibility of death due to shock throughout this monitoring period, and the drugs needed to treat any shock induced by vaccination are costly. Third, blood-borne parasites may be present in the blood vaccine and be transmitted to the vaccinates. Finally, the blood vaccine requires a cold chain to preserve the vaccine.

Clearly, a safer, more effective vaccine that is easily administered would be particularly advantageous. For these reasons, and with the advent of new methods in biotechnology, investigators have concentrated recently on the development of new types of vaccines, including recombinant vaccines. However, recombinant vaccine antigens must be carefully selected and presented to the immune system such that shared epitopes are recognized. These factors have contributed to the search for effective vaccines.

A protective vaccine against rickettsiae that elicits a complete immune response can be advantageous. A few antigens which potentially can be useful as vaccines have now been identified and sequenced for various pathogenic rickettsia. The genes encoding the antigens and that can be employed to recombinantly produce those antigen have also been identified and sequenced. Certain protective antigens identified for *R. rickettsii*, *R. conorii*, and *R. prowazekii* (e.g., rOmpA and rOmpB) are large (>100 kDa), dependent on retention of native conformation for protective efficacy, but are often degraded when produced in recombinant systems. This presents technical and quality-control problems if purified recombinant proteins are to be included in a vaccine. The mode of presentation of a recombinant antigen to the immune system can also be an important factor in the immune response.

Nucleic acid vaccination has been shown to induce protective immune responses in non-viral systems and in diverse animal species (Special Conference Issue, WHO meeting on nucleic acid vaccines [1994] *Vaccine* 12:1491). Nucleic acid vaccination has induced cytotoxic lymphocyte (CTL), T-helper 1, and antibody responses, and has been shown to be protective against disease (Ulmer, J., J. Donnelly, S. Parker *et al.* [1993] *Science* 259:1745). For example, direct intramuscular injection of mice with DNA encoding the influenza nucleoprotein caused the production of high titer antibodies, nucleoprotein-specific CTLs, and protection against viral challenge. Immunization of mice with plasmid DNA encoding the *Plasmodium yoelii* circumsporozoite protein induced high antibody titers against malaria sporozoites and CTLs, and protection against challenge infection (Sedegah, M., R. Hedstrom, P. Hobart, S. Hoffman [1994] *Proc. Natl. Acad. Sci. USA* 91:9866). Cattle immunized with plasmids encoding bovine herpesvirus 1 (BHV-1) glycoprotein IV developed neutralizing antibody and were partially

protected (Cox, G., T. Zamb, L. Babiuk [1993] *J. Virol.* 67:5664). However, it has been a question in the field of immunization whether the recently discovered technology of nucleic acid vaccines can provide improved protection against an antigenic drift variant. Moreover, it has not heretofore been recognized or suggested that nucleic acid vaccines may be successful to protect against rickettsial disease or that a major surface protein conserved in rickettsia was protective against disease.

#### Brief Summary of the Invention

Disclosed and claimed here are novel vaccines for conferring immunity to rickettsia infection, including *Cowdria ruminantium* causing heartwater. Also disclosed are novel nucleic acid compositions and methods of using those compositions, including to confer immunity in a susceptible host. Also disclosed are novel materials and methods for diagnosing infections by *Ehrlichia* in humans or animals.

One aspect of the subject invention concerns a nucleic acid, e.g., DNA or mRNA, vaccine containing the major antigenic protein 1 gene (MAP1) or the major antigenic protein 2 gene (MAP2) of rickettsial pathogens. In one embodiment, the nucleic acid vaccines can be driven by the human cytomegalovirus (HCMV) enhancer-promoter. In studies immunizing mice by intramuscular injection of a DNA vaccine composition according to the subject invention, immunized mice seroconverted and reacted with MAP1 in antigen blots. Splenocytes from immunized mice, but not from control mice immunized with vector only, proliferated in response to recombinant MAP1 and rickettsial antigens in *in vitro* lymphocyte proliferation tests. In experiments testing different DNA vaccine dose regimens, increased survival rates as compared to controls were observed on challenge with rickettsia. Accordingly, the subject invention concerns the discovery that DNA vaccines can induce protective immunity against rickettsial disease or death resulting therefrom.

#### Brief Description of the Drawings

Figures 1A-1C show a comparison of the amino acid sequences from alignment of the three rickettsial proteins, namely, *Cowdria ruminantium* (C.r.), *Ehrlichia chaffeensis* (E.c.), and *Anaplasma marginale* (A.m.).

Figures 2A-2C shows the DNA sequence of the 28 kDa gene locus cloned from *E. chaffeensis* (Fig. 2A-2B) and *E. canis* (Fig. 2C). One letter amino acid codes for the deduced protein sequences are presented below the nucleotide sequence. The proposed sigma-70-like promoter sequences (38) are presented in bold and underlined text as -10 and -35 (consensus -35

and -10 sequences are TTGACA and TATAAT, respectively). Similarly, consensus ribosomal binding sites and transcription terminator sequences (bold letter sequence) are identified. G-rich regions identified in the *E. chaffeensis* sequence are underlined. The conserved sequences from within the coding regions selected for RT-PCR assay are identified with italics and underlined text.

Figure 3A shows the complete sequence of the MAP2 homolog of *Ehrlichia canis*. The arrow (→) represents the predicted start of the mature protein. The asterisk (\*) represents the stop codon. Underlined nucleotides 5' to the open reading frame with -35 and -10 below represent predicted promoter sequences. Double underlined nucleotides represent the predicted ribosomal binding site. Underlined nucleotides 3' to the open reading frame represent possible transcription termination sequences.

Figure 3B shows the complete sequence of the MAP2 homolog of *Ehrlichia chaffeensis*. The arrow (→) represents the predicted start of the mature protein. The asterisk (\*) represents the stop codon. Underlined nucleotides 5' to the open reading frame with -35 and -10 below represent predicted promoter sequences. Double underlined nucleotides represent the predicted ribosomal binding site. Underlined nucleotides 3' to the open reading frame represent possible transcription termination sequences.

#### Brief Description of the Sequences

SEQ ID NO. 1 is the coding sequence of the MAP1 gene from *Cowdria ruminantium* (Highway isolate).

SEQ ID NO. 2 is the polypeptide encoded by the polynucleotide of SEQ ID NO. 1.

SEQ ID NO. 3 is the coding sequence of the MAP1 gene from *Ehrlichia chaffeensis*.

SEQ ID NO. 4 is the polypeptide encoded by the polynucleotide of SEQ ID NO. 3.

SEQ ID NO. 5 is the *Anaplasma marginale* MSP4 gene coding sequence.

SEQ ID NO. 6 is the polypeptide encoded by the polynucleotide of SEQ ID NO. 5.

SEQ ID NO. 7 is a partial coding sequence of the VSA1 gene from *Ehrlichia chaffeensis*, also shown in Figures 2A-2B.

SEQ ID NO. 8 is the coding sequence of the VSA2 gene from *Ehrlichia chaffeensis*, also shown in Figures 2A-2B.

SEQ ID NO. 9 is the coding sequence of the VSA3 gene from *Ehrlichia chaffeensis*, also shown in Figures 2A-2B.

SEQ ID NO. 10 is the coding sequence of the VSA4 gene from *Ehrlichia chaffeensis*, also shown in Figures 2A-2B.

SEQ ID NO. 11 is a partial coding sequence of the VSA5 gene from *Ehrlichia chaffeensis*, also shown in Figures 2A-2B.

SEQ ID NO. 12 is the coding sequence of the VSA1 gene from *Ehrlichia canis*, also shown in Figure 2C.

5 SEQ ID NO. 13 is a partial coding sequence of the VSA2 gene from *Ehrlichia canis*, also shown in Figure 2C.

**SEQ ID NO. 14** is the polypeptide encoded by the polynucleotide of **SEQ ID NO. 7**, also shown in Figures 2A-2B.

SEQ ID NO. 15 is the polypeptide encoded by the polynucleotide of SEQ ID NO. 8,  
10 also shown in Figures 2A-2B.

**SEQ ID NO. 16** is the polypeptide encoded by the polynucleotide of **SEQ ID NO. 9**, also shown in Figures 2A-2B.

**SEQ ID NO. 17** is the polypeptide encoded by the polynucleotide of **SEQ ID NO. 10**, also shown in Figures 2A-2B.

15                    **SEQ ID NO. 18** is the polypeptide encoded by the polynucleotide of **SEQ ID NO. 11**,  
also shown in Figures 2A-2B.

**SEQ ID NO. 19** is the polypeptide encoded by the polynucleotide of **SEQ ID NO. 12**, also shown in Figure 2C.

SEQ ID NO. 20 is the polypeptide encoded by the polynucleotide of SEQ ID NO. 13,  
20 also shown in Figure 2C.

SEQ ID NO. 21 is the coding sequence of the MAP2 gene from *Ehrlichia canis*, also shown in Figure 3A.

**SEQ ID NO. 22** is the coding sequence of the MAP2 gene from *Ehrlichia chaffeensis*, also shown in Figure 3B.

25                    **SEQ ID NO. 23** is the polypeptide encoded by the polynucleotide of SEQ ID NO. 21,  
also shown in Figure 3A.

**SEQ ID NO. 24** is the polypeptide encoded by the polynucleotide of **SEQ ID NO. 22**, also shown in Figure 3B.

30 Detailed Disclosure of the Invention

In one embodiment, the subject invention concerns a novel strategy, termed nucleic acid vaccination, for eliciting an immune response protective against rickettsial disease. The subject invention also concerns novel compositions that can be employed according to this novel strategy for eliciting a protective immune response. According to the subject invention,



recombinant plasmid DNA or mRNA encoding an antigen of interest is inoculated directly into the human or animal host where the antigen is expressed and an immune response induced. Advantageously, problems of protein purification, as can be encountered with antigen delivery using live vectors, can be virtually eliminated by employing the compositions or methods according to the subject invention. Unlike live vector delivery, the subject invention can provide a further advantage in that the DNA or RNA does not replicate in the host, but remains episomal with gene expression directed for as long as 19 months or more post-injection. See, for example, Wolff, J.A., J.J. Ludike, G. Acsadi, P. Williams, A. Jani (1992) *Hum. Mol. Genet.* 1:363. A complete immune response can be obtained as recombinant antigen is synthesized intracellularly and presented to the host immune system in the context of autologous class I and class II MHC molecules.

In one embodiment, the subject invention concerns nucleic acids and compositions comprising those nucleic acids that can be effective in protecting an animal from disease or death caused by rickettsia. For example, a nucleic acid vaccine of the subject invention has been shown to be protective against *Cowdria ruminantium*, the causative agent of heartwater in domestic ruminants. Accordingly, DNA sequences of rickettsial genes, e.g., MAP1 or homologues thereof, can be used as nucleic acid vaccines against human and animal rickettsial diseases. The MAP1 gene used to obtain this protection is also present in other rickettsiae including *Anaplasma marginale*, *Ehrlichia canis*, and in a causative agent of human ehrlichiosis, *Ehrlichia chaffeensis* (van Vliet, A., F. Jongejan, M. van Kleef, B. van der Zeijst [1994] *Infect. Immun.* 62:1451). The MAP1 gene or a MAP1-like gene can also be found in certain *Rickettsia* spp. MAP1-like genes from *Ehrlichia chaffeensis* and *Ehrlichia canis* have now been cloned and sequenced. These MAP-1 homologs are also referred to herein as Variable Surface Antigen (VSA) genes.

The present invention also concerns polynucleotides encoding MAP2 or MAP2 homologs from *Ehrlichia canis* and *Ehrlichia chaffeensis*. MAP2 polynucleotide sequences of the invention can be used as vaccine compositions and in diagnostic assays. The polynucleotides can also be used to produce the MAP2 polypeptides encoded thereby.

Compositions comprising the subject polynucleotides can include appropriate nucleic acid vaccine vectors (plasmids), which are commercially available (e.g., Vical, San Diego, CA). In addition, the compositions can include a pharmaceutically acceptable carrier, e.g., saline. The pharmaceutically acceptable carriers are well known in the art and also are commercially available. For example, such acceptable carriers are described in E.W. Martin's *Remington's Pharmaceutical Science*, Mack Publishing Company, Easton, PA.

The subject invention also concerns polypeptides encoded by the subject polynucleotides. Specifically exemplified are the polypeptides encoded by the MAP-1 and VSA genes of *C. rumimontium*, *E. chaffeensis*, *E. canis* and the MP4 gene of *Anaplasma marginale*. Polypeptides uncoded by *E. chaffeensis* and *E. canis* MAP2 genes are also exemplified herein.

5

Also encompassed within the scope of the present invention are fragments and variants of the exemplified polynucleotides. Variants include polynucleotides and/or polypeptides having base or amino acid additions, deletions and substitutions in the sequence of the subject molecule so long as those variants have substantially the same activity or serologic reactivity as the native molecules. Also included are allelic variants of the subject polynucleotides. The polypeptides and peptides of the present invention can be used to raise antibodies that are reactive with the polypeptides disclosed herein. The polypeptides and peptides can also be used as molecular weight markers.

10

Another aspect of the subject invention concerns antibodies reactive with MAP-1 and MAP2 polypeptides disclosed herein. Antibodies can be monoclonal or polyclonal and can be produced using standard techniques known in the art. Antibodies of the invention can be used in diagnostic and therapeutic applications.

15

In a specific embodiment, the subject invention concerns a DNA vaccine (e.g., VCL1010/MAP1) containing the major antigenic protein 1 gene (MAP1) driven by the human cytomegalovirus (HCMV) enhancer-promoter injected intramuscularly into 8-10 week-old female DBA/2 mice after treating them with 50  $\mu$ l/muscle of 0.5% bupivacaine 3 days previously. Up to 75% of the VCL1010/MAP1-immunized mice seroconverted and reacted with MAP1 in antigen blots. Splenocytes from immunized mice, but not from control mice immunized with VCL1010 DNA (plasmid vector, Vical, San Diego) proliferated in response to recombinant MAP1 and *C. ruminantium* antigens in *in vitro* lymphocyte proliferation tests. These proliferating cells from mice immunized with VCL1010/MAP1 DNA secreted IFN-gamma and IL-2 at concentrations ranging from 610 pg/ml and 152 pg/ml to 1290 pg/ml and 310 pg/ml, respectively. In experiments testing different VCL1010/MAP1 DNA vaccine dose regimens (25-100  $\mu$ g/dose, 2 or 4 immunizations), survival rates of 23% to 88% (35/92 survivors/total in all VCL1010/MAP1 immunized groups) were observed on challenge with 30LD50 of *C. ruminantium*. Survival rates of 0% to 3% (1/144 survivors/total in all control groups) were recorded for control mice immunized similarly with VCL1010 DNA or saline. Accordingly, the subject invention concerns the discovery that the gene encoding the MAP1 protein can induce protective immunity as a DNA vaccine against rickettsial disease.

20

25

30

The nucleic acid sequences described herein have other uses as well. For example, the nucleic acids of the subject invention can be useful as probes to identify complementary sequences within other nucleic acid molecules or genomes. Such use of probes can be applied to identify or distinguish infectious strains of organisms in diagnostic procedures or in rickettsial research where identification of particular organisms or strains is needed. As is well known in the art, probes can be made by labeling the nucleic acid sequences of interest according to accepted nucleic acid labeling procedures and techniques. A person of ordinary skill in the art would recognize that variations or fragments of the disclosed sequences which can specifically and selectively hybridize to the DNA of rickettsia can also function as a probe. It is within the ordinary skill of persons in the art, and does not require undue experimentation in view of the description provided herein, to determine whether a segment of the claimed DNA sequences is a fragment or variant which has characteristics of the full sequence, *e.g.*, whether it specifically and selectively hybridizes or can confer protection against rickettsial infection in accordance with the subject invention. In addition, with the benefit of the subject disclosure describing the specific sequences, it is within the ordinary skill of those persons in the art to label hybridizing sequences to produce a probe.

It is also well known in the art that restriction enzymes can be used to obtain functional fragments of the subject DNA sequences. For example, *Bal31* exonuclease can be conveniently used for time-controlled limited digestion of DNA (commonly referred to as "erase-a-base" procedures). See, for example, Maniatis *et al.* (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York; Wei *et al.* (1983) *J. Biol. Chem.* 258:13006-13512.

In addition, the nucleic acid sequences of the subject invention can be used as molecular weight markers in nucleic acid analysis procedures.

Following are examples which illustrate procedures for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

#### Example 1

A nucleic acid vaccine construct was tested in animals for its ability to protect against death caused by infection with the rickettsia *Cowdria ruminantium*. The vaccine construct tested was the MAP1 gene of *C. ruminantium* inserted into plasmid VCL1010 (Vical, San Diego) under control of the human cytomegalovirus promoter-enhancer and intron A. In this study, seven groups containing 10 mice each were injected twice at 2-week intervals with either 100, 75, 50,

or 25 µg VCL1010/MAP1 DNA (V/M in Table 1 below), or 100, 50 µg VCL1010 DNA (V in Table 1) or saline (Sal.), respectively. Two weeks after the last injections, 8 mice/group were challenged with 30LD50 of *C. ruminantium* and clinical symptoms and survival monitored. The remaining 2 mice/group were not challenged and were used for lymphocyte proliferation tests and cytokine measurements. The results of the study are summarized in Table 1, below:

Table 1							
	100 µg V/M	75 µg V/M	50 µg V/M	25 µg V/M	100 µg V	50 µg V	Sal.
Survived	5	7	5	3	0	0	0
Died	3	1	3	5	8	8	8

The VCL1010/MAP1 nucleic acid vaccine increased survival on challenge in all groups, with a total of 20/30 mice surviving compared to 0/24 in the control groups.

This study was repeated with another 6 groups, each containing 33 mice (a total of 198 mice). Three groups received 75 µg VCL1010/MAP1 DNA or VCL1010 DNA or saline (4 injections in all cases). Two weeks after the last injection, 30 mice/group were challenged with 30LD50 of *C. ruminantium* and 3 mice/group were sacrificed for lymphocyte proliferation tests and cytokine measurements. The results of this study are summarized in Table 2, below:

Table 2						
	V/M 2 inj.	V 2 inj.	Sal. 2 inj.	V/M 4 inj.	V 4 inj.	Sal. 4 inj.
Survived	7	0	0	8	0	1
Died*	23	30	30	22	30	29

\*In mice that died in both V/M groups, there was an increase in mean survival time of approximately 4 days compared to the controls ( $p < 0.05$ ).

Again, as summarized in Table 2, the VCL1010/MAP1 DNA vaccine increased the numbers of mice surviving in both immunized groups, although there was no apparent benefit of 2 additional injections. In these two experiments, there were a cumulative total of 35/92 (38%) surviving mice in groups receiving the VCL1010/MAP1 DNA vaccine compared to 1/144 (0.7%) surviving mice in the control groups. In both immunization and challenge trials

described above, splenocytes from VCL1010/MAP1 immunized mice, but not from control mice, specifically proliferated to recombinant MAP1 protein and to *C. ruminantium* in lymphocyte proliferation tests. These proliferating splenocytes secreted IL-2 and gamma-interferon at concentrations up to 310 and 1290 pg/ml respectively. These data show that protection against rickettsial infections can be achieved with a DNA vaccine. In addition, these experiments show MAP1-related proteins as vaccine targets.

### Example 2

The MAP1 protein of *C. ruminantium* has significant similarity to MSP4 of *A. marginale*, and related molecules may also be presenting other rickettsial pathogens. To prove this, we used primers based on regions conserved between *C. ruminantium* and *A. marginale* in PCR to clone a MAP1-like gene from *E. chaffeensis*. The amino acid sequence derived from the cloned *E. chaffeensis* MAP1-like gene, and alignment with the corresponding genes of *C. ruminantium* and *A. marginale* is shown in Figure 1. We have now identified the regions of MAP1-like genes which are highly conserved between *Ehrlichia*, *Cowdria*, and *Anaplasma* and which can allow cloning of the analogous genes from other rickettsiae.

### Example 3 – Cloning and sequence analysis of MAP1 homologue genes of *E. chaffeensis* and *E. canis*

Genes homologous to the major surface protein of *C. ruminantium* MAP1 were cloned from *E. chaffeensis* and *E. canis* by using PCR cloning strategies. The cloned segments represent a 4.6 kb genomic locus of *E. chaffeensis* and a 1.6 kb locus of *E. canis*. DNA sequence generated from these clones was assembled and is presented along with the deduced amino acid sequence in Figures 2A-2B (SEQ ID NOs. 7-11 and 14-18) and Figure 2C (SEQ ID NOs. 12-13 and 19-20). Significant features of the DNA include five very similar but nonidentical open reading frames (ORFs) for *E. chaffeensis* and two very similar, nonidentical ORFs for the *E. canis* cloned locus. The ORFs for both *Ehrlichia* spp. are separated by noncoding sequences ranging from 264 to 310 base pairs. The noncoding sequences have a higher A+T content (71.6% for *E. chaffeensis* and 76.1% for *E. canis*) than do the coding sequences (63.5% for *E. chaffeensis* and 68.0% for *E. canis*). A G-rich region -200 bases upstream from the initiation codon, sigma-70-like promoter sequences, putative ribosome binding sites (RBS), termination codons, and palindromic sequences near the termination codons are found in each of the *E. chaffeensis* noncoding sequences. The *E. canis* noncoding sequence has the same feature except for the G-rich region (Figure 2C; SEQ ID NOs. 12-13 and 19-20).

Sequence comparisons of the ORFs at the nucleotide and translated amino acid levels revealed a high degree of similarity between them. The similarity spanned the entire coding sequences, except in three regions where notable sequence variations were observed including some deletions/insertions (Variable Regions I, II and III). Despite the similarities, no two ORFs are identical. The cloned ORF 2, 3 and 4 of *E. chaffeensis* have complete coding sequences. The ORF1 is a partial gene having only 143 amino acids at the C-terminus whereas the ORF5 is nearly complete but lacks 5-7 amino acids and a termination codon. The cloned ORF2 of *E. canis* also is a partial gene lacking a part of the C-terminal sequence. The overall similarity between different ORFs at the amino acid level is 56.0% to 85.4% for *E. chaffeensis*, whereas for *E. canis* it is 53.3%. The similarity of *E. chaffeensis* ORFs to the MAP1 coding sequences reported for *C. ruminantium* isolates ranged from 55.5% to 66.7%, while for *E. canis* to *C. ruminantium* it is 48.5% to 54.2%. Due to their high degree of similarity to MAP1 surface antigen genes of *C. ruminantium* and since they are nonidentical to each other, the *E. chaffeensis* and *E. canis* ORFs are referred to herein as putative Variable Surface Antigen (VSA) genes. The apparent molecular masses of the predicted mature proteins of *E. chaffeensis* were 28.75 kDa for VSA2, 27.78 for VSA3, and 27.95 for VSA4, while *E. canis* VSA1 was slightly higher at 29.03 kDa. The first 25 amino acids in each VSA coding sequence were eliminated when calculating the protein size since they markedly resembled the signal sequence of *C. ruminantium* MAP1 and presumably would be absent from the mature protein. Predicted protein sizes for *E. chaffeensis* VSA1 and VSA5, and *E. canis* VSA2 were not calculated since the complete genes were not cloned.

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and the scope of the appended claims.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

Applicant Name(s): University of Florida  
Street Address: 223 Grinter Hall  
City: Gainesville  
State/Province: Florida  
Country: US  
Postal Code/Zip: 32611  
Phone number: (352) 392-8929 Fax: (352) 392-6600

(ii) TITLE OF INVENTION: Nucleic Acid Vaccines Against  
Rickettsial Diseases and Methods of Use

(iii) NUMBER OF SEQUENCES: 24

## (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
(B) STREET: 2421 N.W. 41st Street, Suite A-1  
(C) CITY: Gainesville  
(D) STATE: FL  
(E) COUNTRY: USA  
(F) ZIP: 32606

## (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

## (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: PCT  
(B) FILING DATE: 17 October 1997  
(C) CLASSIFICATION:

## (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Pace, Doran R.  
(B) REGISTRATION NUMBER: 38,261  
(C) REFERENCE/DOCKET NUMBER: UF-167C1

## (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 352-375-8100  
(B) TELEFAX: 352-372-5800

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 864 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..861

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG AAT TGC AAG AAA ATT TTT ATC ACA AGT ACA CTA ATA TCA TTA GTG	48
Met Asn Cys Lys Lys Ile Phe Ile Thr Ser Thr Leu Ile Ser Leu Val	
1 5 10 15	
TCA TTT TTA CCT GGT GTG TCC TTT TCT GAT GTA ATA CAG GAA GAC AGC	96
Ser Phe Leu Pro Gly Val Ser Phe Ser Asp Val Ile Gln Glu Asp Ser	
20 25 30	
AAC CCA GCA GGC AGT GTT TAC ATT AGC GCA AAA TAC ATG CCA ACT GCA	144
Asn Pro Ala Gly Ser Val Tyr Ile Ser Ala Lys Tyr Met Pro Thr Ala	
35 40 45	
TCA CAT TTT GGT AAA ATG TCA ATC AAA GAA GAT TCA AAA AAT ACT CAA	192
Ser His Phe Gly Lys Met Ser Ile Lys Glu Asp Ser Lys Asn Thr Gln	
50 55 60	
ACG GTA TTT GGT CTA AAA AAA GAT TGG GAT GGC GTT AAA ACA CCA TCA	240
Thr Val Phe Gly Leu Lys Lys Asp Trp Asp Gly Val Lys Thr Pro Ser	
65 70 75 80	
GAT TCT AGC AAT ACT AAT TCT ACA ATT TTT ACT GAA AAA GAC TAT TCT	288
Asp Ser Ser Asn Thr Asn Ser Thr Ile Phe Thr Glu Lys Asp Tyr Ser	
85 90 95	
TTC AGA TAT GAA AAC AAT CCG TTT TTA GGT TTC GCT GGA GCA ATT GGG	336
Phe Arg Tyr Glu Asn Asn Pro Phe Leu Gly Phe Ala Gly Ala Ile Gly	
100 105 110	
TAC TCA ATG AAT GGA CCA AGA ATA GAG TTC GAA GTA TCC TAT GAA ACT	384
Tyr Ser Met Asn Gly Pro Arg Ile Glu Phe Glu Val Ser Tyr Glu Thr	
115 120 125	
TTT GAT GTA AAA AAC CTA GGT GGC AAC TAT AAA AAC AAC GCA CAC ATG	432
Phe Asp Val Lys Asn Leu Gly Gly Asn Tyr Lys Asn Asn Ala His Met	
130 135 140	
TAC TGT GCT TTA GAT ACA GCA GCA CAA AAT AGC ACT AAT GGC GCA GGA	480
Tyr Cys Ala Leu Asp Thr Ala Ala Gln Asn Ser Thr Asn Gly Ala Gly	
145 150 155 160	
TTA ACT ACA TCT GTT ATG GTA AAA AAC GAA AAT TTA ACA AAT ATA TCA	528
Leu Thr Thr Ser Val Met Val Lys Asn Glu Asn Leu Thr Asn Ile Ser	
165 170 175	
TTA ATG TTA AAT GCG TGT TAT GAT ATC ATG CTT GAT GGA ATA CCA GTT	576
Leu Met Leu Asn Ala Cys Tyr Asp Ile Met Leu Asp Gly Ile Pro Val	
180 185 190	



TCT CCA TAT GTA TGT GCA GGT ATT GGC ACT GAC TTA GTG TCA GTA ATT	624
Ser Pro Tyr Val Cys Ala Gly Ile Gly Thr Asp Leu Val Ser Val Ile	
195 200 205	
AAT GCT ACA AAT CCT AAA TTA TCT TAT CAA GGA AAG CTA GGC ATA AGT	672
Asn Ala Thr Asn Pro Lys Leu Ser Tyr Gln Gly Lys Leu Gly Ile Ser	
210 215 220	
TAC TCA ATC AAT TCT GAA GCT TCT ATC TTT ATC GGT GGA CAT TTC CAT	720
Tyr Ser Ile Asn Ser Glu Ala Ser Ile Phe Ile Gly Gly His Phe His	
225 230 235 240	
AGA GTT ATA GGT AAT GAA TTT AAA GAT ATT GCT ACC TTA AAA ATA TTT	768
Arg Val Ile Gly Asn Glu Phe Lys Asp Ile Ala Thr Leu Lys Ile Phe	
245 250 255	
ACT TCA AAA ACA GGA ATA TCT AAT CCT GGC TTT GCA TCA GCA ACA CTT	816
Thr Ser Lys Thr Gly Ile Ser Asn Pro Gly Phe Ala Ser Ala Thr Leu	
260 265 270	
GAT GTT TGT CAC TTT GGT ATA GAA ATT GGA GGA AGG TTT GTA TTT	861
Asp Val Cys His Phe Gly Ile Glu Ile Gly Gly Arg Phe Val Phe	
275 280 285	
TAA	864

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 287 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asn Cys Lys Lys Ile Phe Ile Thr Ser Thr Leu Ile Ser Leu Val	
1 5 10 15	
Ser Phe Leu Pro Gly Val Ser Phe Ser Asp Val Ile Gln Glu Asp Ser	
20 25 30	
Asn Pro Ala Gly Ser Val Tyr Ile Ser Ala Lys Tyr Met Pro Thr Ala	
35 40 45	
Ser His Phe Gly Lys Met Ser Ile Lys Glu Asp Ser Lys Asn Thr Gln	
50 55 60	
Thr Val Phe Gly Leu Lys Lys Asp Trp Asp Gly Val Lys Thr Pro Ser	
65 70 75 80	
Asp Ser Ser Asn Thr Asn Ser Thr Ile Phe Thr Glu Lys Asp Tyr Ser	
85 90 95	

Phe Arg Tyr Glu Asn Asn Pro Phe Leu Gly Phe Ala Gly Ala Ile Gly  
 100 105 110  
 Tyr Ser Met Asn Gly Pro Arg Ile Glu Phe Glu Val Ser Tyr Glu Thr  
 115 120 125  
 Phe Asp Val Lys Asn Leu Gly Gly Asn Tyr Lys Asn Asn Ala His Met  
 130 135 140  
 Tyr Cys Ala Leu Asp Thr Ala Ala Gln Asn Ser Thr Asn Gly Ala Gly  
 145 150 155 160  
 Leu Thr Thr Ser Val Met Val Lys Asn Glu Asn Leu Thr Asn Ile Ser  
 165 170 175  
 Leu Met Leu Asn Ala Cys Tyr Asp Ile Met Leu Asp Gly Ile Pro Val  
 180 185 190  
 Ser Pro Tyr Val Cys Ala Gly Ile Gly Thr Asp Leu Val Ser Val Ile  
 195 200 205  
 Asn Ala Thr Asn Pro Lys Leu Ser Tyr Gln Gly Lys Leu Gly Ile Ser  
 210 215 220  
 Tyr Ser Ile Asn Ser Glu Ala Ser Ile Phe Ile Gly Gly His Phe His  
 225 230 235 240  
 Arg Val Ile Gly Asn Glu Phe Lys Asp Ile Ala Thr Leu Lys Ile Phe  
 245 250 255  
 Thr Ser Lys Thr Gly Ile Ser Asn Pro Gly Phe Ala Ser Ala Thr Leu  
 260 265 270  
 Asp Val Cys His Phe Gly Ile Glu Ile Gly Gly Arg Phe Val Phe  
 275 280 285

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 842 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..840

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG AAT TAC AAA AAA AGT TTC ATA ACA GCG ATT GAT ATC ATT AAT ATC Met Asn Tyr Lys Lys Ser Phe Ile Thr Ala Ile Asp Ile Ile Asn Ile 290 295 300	48
CTT CTC TTA CCT GGA GTA TCA TTT TCC GAC CCA AGG CAG GTA GTG GTC Leu Leu Leu Pro Gly Val Ser Phe Ser Asp Pro Arg Gln Val Val Val 305 310 315	96
ATT AAC GGT AAT TTC TAC ATC AGT GGA AAA TAC GAT GCC AAG GCT TCG Ile Asn Gly Asn Phe Tyr Ile Ser Gly Lys Tyr Asp Ala Lys Ala Ser 320 325 330 335	144
CAT TTT GGA GTA TTC TCT GCT AAG GAA GAA AGA AAT ACA ACA GTT GGA His Phe Gly Val Phe Ser Ala Lys Glu Glu Arg Asn Thr Thr Val Gly 340 345 350	192
GTG TTT GGA CTG AAG CAA AAT TGG GAC GGA AGC GCA ATA TCC AAC TCC Val Phe Gly Leu Lys Gln Asn Trp Asp Gly Ser Ala Ile Ser Asn Ser 355 360 365	240
TCC CCA AAC GAT GTA TTC ACT GTC TCA AAT TAT TCA TTT AAA TAT GAA Ser Pro Asn Asp Val Phe Thr Val Ser Asn Tyr Ser Phe Lys Tyr Glu 370 375 380	288
AAC AAC CCG TTT TTA GGT TTT GCA GGA GCT ATT GGT TAC TCA ATG GAT Asn Asn Pro Phe Leu Gly Phe Ala Gly Ala Ile Gly Tyr Ser Met Asp 385 390 395	336
GGT CCA AGA ATA GAG CTT GAA GTA TCT TAT GAA ACA TTT GAT GTA AAA Gly Pro Arg Ile Glu Leu Glu Val Ser Tyr Glu Thr Phe Asp Val Lys 400 405 410 415	384
AAT CAA GGT AAC AAT TAT AAG AAT GAA GCA CAT AGA TAT TGT GCT CTA Asn Gln Gly Asn Asn Tyr Lys Asn Glu Ala His Arg Tyr Cys Ala Leu 420 425 430	432
TCC CAT AAC TCA GCA GCA GAC ATG AGT AGT GCA AGT AAT AAT TTT GTC Ser His Asn Ser Ala Ala Asp Met Ser Ser Ala Ser Asn Asn Phe Val 435 440 445	480
TTT CTA AAA AAT GAA GGA TTA CTT GAC ATA TCA TTT ATG CTG AAC GCA Phe Leu Lys Asn Glu Gly Leu Leu Asp Ile Ser Phe Met Leu Asn Ala 450 455 460	528
TGC TAT GAC GTA GTA GGC GAA GGC ATA CCT TTT TCT CCT TAT ATA TGC Cys Tyr Asp Val Val Gly Glu Gly Ile Pro Phe Ser Pro Tyr Ile Cys 465 470 475	576
GCA GGT ATC GGT ACT GAT TTA GTA TCC ATG TTT GAA GCT ACA AAT CCT Ala Gly Ile Gly Thr Asp Leu Val Ser Met Phe Glu Ala Thr Asn Pro 480 485 490 495	624
AAA ATT TCT TAC CAA GGA AAG TTA GGT TTA AGC TAC TCT ATA AGC CCA Lys Ile Ser Tyr Gln Gly Lys Leu Gly Leu Ser Tyr Ser Ile Ser Pro 500 505 510	672

GAA GCT TCT GTG TTT ATT GGT GGG CAC TTT CAT AAG GTA ATA GGG AAC	720
Glu Ala Ser Val Phe Ile Gly Gly His Phe His Lys Val Ile Gly Asn	
515 520 525	
GAA TTT AGA GAT ATT CCT ACT ATA ATA CCT ACT GGA TCA ACA CTT GCA	768
Glu Phe Arg Asp Ile Pro Thr Ile Ile Pro Thr Gly Ser Thr Leu Ala	
530 535 540	
GGA AAA GGA AAC TAC CCT GCA ATA GTA ATA CTG GAT GTA TGC CAC TTT	816
Gly Lys Gly Asn Tyr Pro Ala Ile Val Ile Leu Asp Val Cys His Phe	
545 550 555	
GGA ATA GAA ATG GGA GGA AGG TTT AA	842
Gly Ile Glu Met Gly Gly Arg Phe	
560 565	

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 280 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asn Tyr Lys Lys Ser Phe Ile Thr Ala Ile Asp Ile Ile Asn Ile	
1 5 10 15	
Leu Leu Leu Pro Gly Val Ser Phe Ser Asp Pro Arg Gln Val Val Val	
20 25 30	
Ile Asn Gly Asn Phe Tyr Ile Ser Gly Lys Tyr Asp Ala Lys Ala Ser	
35 40 45	
His Phe Gly Val Phe Ser Ala Lys Glu Glu Arg Asn Thr Thr Val Gly	
50 55 60	
Val Phe Gly Leu Lys Gln Asn Trp Asp Gly Ser Ala Ile Ser Asn Ser	
65 70 75 80	
Ser Pro Asn Asp Val Phe Thr Val Ser Asn Tyr Ser Phe Lys Tyr Glu	
85 90 95	
Asn Asn Pro Phe Leu Gly Phe Ala Gly Ala Ile Gly Tyr Ser Met Asp	
100 105 110	
Gly Pro Arg Ile Glu Leu Glu Val Ser Tyr Glu Thr Phe Asp Val Lys	
115 120 125	
Asn Gln Gly Asn Asn Tyr Lys Asn Glu Ala His Arg Tyr Cys Ala Leu	
130 135 140	

Ser His Asn Ser Ala Ala Asp Met Ser Ser Ala Ser Asn Asn Phe Val  
 145 150 155 160

Phe Leu Lys Asn Glu Gly Leu Leu Asp Ile Ser Phe Met Leu Asn Ala  
 165 170 175

Cys Tyr Asp Val Val Gly Glu Gly Ile Pro Phe Ser Pro Tyr Ile Cys  
 180 185 190

Ala Gly Ile Gly Thr Asp Leu Val Ser Met Phe Glu Ala Thr Asn Pro  
 195 200 205

Lys Ile Ser Tyr Gln Gly Lys Leu Gly Leu Ser Tyr Ser Ile Ser Pro  
 210 215 220

Glu Ala Ser Val Phe Ile Gly Gly His Phe His Lys Val Ile Gly Asn  
 225 230 235 240

Glu Phe Arg Asp Ile Pro Thr Ile Ile Pro Thr Gly Ser Thr Leu Ala  
 245 250 255

Gly Lys Gly Asn Tyr Pro Ala Ile Val Ile Leu Asp Val Cys His Phe  
 260 265 270

Gly Ile Glu Met Gly Gly Arg Phe  
 275 280

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 849 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..846

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATG AAT TAC AGA GAA TTG TTT ACA GGG GGC CTG TCA GCA GCC ACA GTC	48
Met Asn Tyr Arg Glu Leu Phe Thr Gly Gly Leu Ser Ala Ala Thr Val	
285 290 295	
TGC GCC TGC TCC CTA CTT GTT AGT GGG GCC GTA GTG GCA TCT CCC ATG	96
Cys Ala Cys Ser Leu Leu Val Ser Gly Ala Val Val Ala Ser Pro Met	
300 305 310	
AGT CAC GAA GTG GCT TCT GAA GGG GGA GTA ATG GGA GGT AGC TTT TAC	144
Ser His Glu Val Ala Ser Glu Gly Gly Val Met Gly Gly Ser Phe Tyr	
315 320 325	

GTG GGT GCG GCC TAC AGC CCA GCA TTT CCT TCT GTT ACC TCG TTC GAC Val Gly Ala Ala Tyr Ser Pro Ala Phe Pro Ser Val Thr Ser Phe Asp 330 335 340	192
ATG CGT GAG TCA AGC AAA GAG ACC TCA TAC GTT AGA GGC TAT GAC AAG Met Arg Glu Ser Ser Lys Glu Thr Ser Tyr Val Arg Gly Tyr Asp Lys 345 350 355 360	240
AGC ATT GCA ACG ATT GAT GTG AGT GTG CCA GCA AAC TTT TCC AAA TCT Ser Ile Ala Thr Ile Asp Val Ser Val Pro Ala Asn Phe Ser Lys Ser 365 370 375	288
GGC TAC ACT TTT GCC TTC TCT AAA AAC TTA ATC ACG TCT TTC GAC GGC Gly Tyr Thr Phe Ala Phe Ser Lys Asn Leu Ile Thr Ser Phe Asp Gly 380 385 390	336
GCT GTG GGA TAT TCT CTG GGA GGA GCC AGA GTG GAA TTG GAA GCG AGC Ala Val Gly Tyr Ser Leu Gly Gly Ala Arg Val Glu Leu Glu Ala Ser 395 400 405	384
TAC AGA AGG TTT GCT ACT TTG GCG GAC GGG CAG TAC GCA AAA AGT GGT Tyr Arg Arg Phe Ala Thr Leu Ala Asp Gly Gln Tyr Ala Lys Ser Gly 410 415 420	432
GCG GAA TCT CTG GCA GCT ATT ACC CGC GAC GCT AAC ATT ACT GAG ACC Ala Glu Ser Leu Ala Ala Ile Thr Arg Asp Ala Asn Ile Thr Glu Thr 425 430 435 440	480
AAT TAC TTC GTA GTC AAA ATT GAT GAA ATC ACA AAC ACC TCA GTC ATG Asn Tyr Phe Val Val Lys Ile Asp Glu Ile Thr Asn Thr Ser Val Met 445 450 455	528
TTA AAT GGC TGC TAT GAC GTG CTG CAC ACA GAT TTA CCT GTG TCC CCG Leu Asn Gly Cys Tyr Asp Val Leu His Thr Asp Leu Pro Val Ser Pro 460 465 470	576
TAT GTA TGT GCC GGG ATA GGC GCA AGC TTT GTT GAC ATC TCT AAG CAA Tyr Val Cys Ala Gly Ile Gly Ala Ser Phe Val Asp Ile Ser Lys Gln 475 480 485	624
GTA ACC ACA AAG CTG GCC TAC AGG GGC AAG GTT GGG ATT AGC TAC CAG Val Thr Thr Lys Leu Ala Tyr Arg Gly Lys Val Gly Ile Ser Tyr Gln 490 495 500	672
TTT ACT CCG GAA ATA TCC TTG GTG GCA GGT GGG TTC TAC CAC GGG CTA Phe Thr Pro Glu Ile Ser Leu Val Ala Gly Gly Phe Tyr His Gly Leu 505 510 515 520	720
TTT GAT GAG TCT TAC AAG GAC ATT CCC GCA CAC AAC AGT GTA AAG TTC Phe Asp Glu Ser Tyr Lys Asp Ile Pro Ala His Asn Ser Val Lys Phe 525 530 535	768
TCT GGA GAA GCA AAA GCC TCA GTC AAA GCG CAT ATT GCT GAC TAC GGC Ser Gly Glu Ala Lys Ala Ser Val Lys Ala His Ile Ala Asp Tyr Gly 540 545 550	816

TTT AAC CTT GGA GCA AGA TTC CTG TTC AGC TAA  
Phe Asn Leu Gly Ala Arg Phe Leu Phe Ser  
555 560

849

## (2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 282 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asn Tyr Arg Glu Leu Phe Thr Gly Gly Leu Ser Ala Ala Thr Val  
1 5 10 15

Cys Ala Cys Ser Leu Leu Val Ser Gly Ala Val Val Ala Ser Pro Met  
20 25 30

Ser His Glu Val Ala Ser Glu Gly Gly Val Met Gly Gly Ser Phe Tyr  
35 40 45

Val Gly Ala Ala Tyr Ser Pro Ala Phe Pro Ser Val Thr Ser Phe Asp  
50 55 60

Met Arg Glu Ser Ser Lys Glu Thr Ser Tyr Val Arg Gly Tyr Asp Lys  
65 70 75 80

Ser Ile Ala Thr Ile Asp Val Ser Val Pro Ala Asn Phe Ser Lys Ser  
85 90 95

Gly Tyr Thr Phe Ala Phe Ser Lys Asn Leu Ile Thr Ser Phe Asp Gly  
100 105 110

Ala Val Gly Tyr Ser Leu Gly Gly Ala Arg Val Glu Leu Glu Ala Ser  
115 120 125

Tyr Arg Arg Phe Ala Thr Leu Ala Asp Gly Gln Tyr Ala Lys Ser Gly  
130 135 140

Ala Glu Ser Leu Ala Ala Ile Thr Arg Asp Ala Asn Ile Thr Glu Thr  
145 150 155 160

Asn Tyr Phe Val Val Lys Ile Asp Glu Ile Thr Asn Thr Ser Val Met  
165 170 175

Leu Asn Gly Cys Tyr Asp Val Leu His Thr Asp Leu Pro Val Ser Pro  
180 185 190

Tyr Val Cys Ala Gly Ile Gly Ala Ser Phe Val Asp Ile Ser Lys Gln  
195 200 205

Val Thr Thr Lys Leu Ala Tyr Arg Gly Lys Val Gly Ile Ser Tyr Gln  
210 215 220

Phe Thr Pro Glu Ile Ser Leu Val Ala Gly Gly Phe Tyr His Gly Leu  
225 230 235 240

Phe Asp Glu Ser Tyr Lys Asp Ile Pro Ala His Asn Ser Val Lys Phe  
245 250 255

Ser Gly Glu Ala Lys Ala Ser Val Lys Ala His Ile Ala Asp Tyr Gly  
260 265 270

Phe Asn Leu Gly Ala Arg Phe Leu Phe Ser  
275 280



Claims

- 1           1. A composition comprising a polynucleotide which encodes a polypeptide having the  
2           characteristic of eliciting an immune response protective against disease or death caused by a  
3           rickettsial pathogen.
- 1           2. The composition, according to claim 1, wherein said rickettsial pathogen is selected  
2           from the group consisting of *Rickettsia* spp., *Ehrlichia* spp., *Anaplasma* spp., and *Cowdria* spp.
- 1           3. The composition, according to claim 1, wherein said polypeptide has an amino acid  
2           sequence selected from the group consisting of SEQ ID NO. 2, SEQ ID NO. 4, SEQ ID NO. 6,  
3           SEQ ID NO. 14, SEQ ID NO. 15, SEQ ID NOS. 16-20, SEQ ID NO. 23, and SEQ ID NO. 24,  
4           or a fragment thereof.
- 1           4. The composition, according to claim 1, wherein said polynucleotide has a nucleic  
2           acid sequence selected from the group consisting of SEQ ID NO. 1, SEQ ID NO. 3, SEQ ID NO.  
3           5, SEQ ID NO. 7, SEQ ID NO. 8, SEQ ID NOS. 9-13, SEQ ID NO. 21, and SEQ ID NO. 22,  
4           or a fragment thereof.
- 1           5. The composition, according to claim 4, wherein said polynucleotide has a nucleic  
2           acid sequence of SEQ ID NO. 3, or a fragment thereof.
- 1           6. The composition, according to claim 1, wherein said polynucleotide further  
2           comprises a nucleic acid vaccine vector.
- 1           7. The composition, according to claim 1, further comprising a pharmaceutically  
2           acceptable carrier.
- 1           8. A polynucleotide encoding a polypeptide having an amino acid sequence selected  
2           from the group consisting of SEQ ID NO. 4, SEQ ID NOS. 14-20, SEQ ID NOS. 23-24, and  
3           fragments thereof.

1           9. The polynucleotide, according to claim 8, said polynucleotide having a nucleic acid  
2           sequence selected from the group consisting of SEQ ID NO. 3, SEQ ID NOS. 7-13, and SEQ  
3           ID NOS. 21-22.

1           10. A method for protecting a susceptible animal host against disease or death caused  
2           by a rickettsial pathogen, said method comprising administering an effective amount of a  
3           polynucleotide encoding polypeptide having the characteristic of eliciting an immune response  
4           protective against said rickettsial pathogen.

1           11. The method, according to claim 10, wherein said rickettsial pathogen is selected  
2           from the group consisting of *Rickettsia* spp., *Ehrlichia* spp., *Anaplasma* spp., and *Cowdria* spp.

1           12. The method, according to claim 10, wherein said polypeptide has an amino acid  
2           sequence selected from the group consisting of SEQ ID NO. 2, SEQ ID NO. 4, SEQ ID NO. 6,  
3           SEQ ID NO. 14, SEQ ID NO. 15, SEQ ID NOS. 16-20, SEQ ID NO. 23, and SEQ ID NO. 24,  
4           or a fragment thereof.

1           13. The method, according to claim 10, wherein said polynucleotide has a nucleic acid  
2           sequence selected from the group consisting of SEQ ID NO. 1, SEQ ID NO. 3, SEQ ID NO. 5,  
3           SEQ ID NO. 7, SEQ ID NO. 8, SEQ ID NOS. 9-13, SEQ ID NO. 21, and SEQ ID NO. 22.

1           14. The method, according to claim 13, wherein said polynucleotide has the nucleic acid  
2           sequence of SEQ ID NO. 1.

1           15. The method, according to claim 13, wherein said polynucleotide has the nucleic acid  
2           sequence of SEQ ID NO. 3.

1           16. The method, according to claim 13, wherein said polynucleotide has the nucleic acid  
2           sequence of SEQ ID NO. 5.

1           17. The method, according to claim 10, wherein said nucleic acid further comprises an  
2           appropriate nucleic acid vector.

1           18. The method, according to claim 10, wherein said composition further comprises a  
2           pharmaceutically acceptable carrier.

1           19. A method for detecting, in a human or animal, antibodies associated with infection  
2           by *Ehrlichia*, wherein said method comprises contacting a biological fluid from said human or  
3           animal with a polypeptide selected from the group consisting of SEQ ID NO. 4, SEQ ID NOS.  
4           14-20, SEQ ID NOS. 23-24, and fragments thereof.

## FIG. 1A

<i>C.r.</i>	ATGAATTGCAAGAAATTTTA-----TCACAAGTACACTAATATCATTAGTG
<i>E.c.</i>	ATGAATTACAAAAAAGTTTCA-----TAACAGCG-ATTGATATCATTAATA
<i>A.m.</i>	ATGAATTACAGAGAAATTGTTTACAGGGGGCCTG-TCAGCAGCC-ACAGTCTGCGCCTGCT
	***** ** ** * * * * *
<i>C.r.</i>	TCATTTT--TACCTGGTGTGTCCTTTTCTGATGTAATACAGGAAGACAGCAACCCAGCAG
<i>E.c.</i>	TCCTTCTCTTACCTGGAGTATCATTTTCCGACCCAGGAGGTAGTGTCA---TTAACG
<i>A.m.</i>	CCCTACTTGTAGTGGGCCGTAGTGGCATCTCCCATGAGTCACGAAGTGGCTTCTGAAG
	* * * * * ** * * * * *
<i>C.r.</i>	GCAGTGTTTACATTAGCGCAAAATACATGCCAACTGCATCACATTTTGGTAAATGTCAA
<i>E.c.</i>	GTAATTTCTACATCAGTGGAAAAATACGATGCCAAGGCTTCGCATTTTGGAGTATTCTCTG
<i>A.m.</i>	GGGGAGTAATGGGAGGTAGCTTTTACGTGGGTGCGGCT-ACAGCCCAGCATTTCCCTTCT
	* * * * * ** * * * * *
<i>C.r.</i>	TCAAAGAAGATTCAAAAAATACTCAAACGGTATTTGGTCTAAAAAAGATTGGGATGGCG
<i>E.c.</i>	CTAAGGAAGAAAGAAATACAACAGTTGGAGTGTTTGGACTGAAGCAAAATTTGGACGGAA
<i>A.m.</i>	GTACCTCGTTTCGACATGCGTGAGTCAAGCAAAAGAGACCTCA--TACGTTAGAGGCTATG
	* * * * * * * * * * *
<i>C.r.</i>	TTAAACACCATCAGATTCTAGCAATACTAATCTACAATTTTACTGAAAAAGACTATT
<i>E.c.</i>	GGCAATATC--CAACTCCTCCCAACGA-----TGATTCACTGTCTCAAAATTATT
<i>A.m.</i>	ACAAGAGCATTGCAACGATTGATGTGAGTGTGCCAGCAAACTTTTCCAAATCTGGCTACA
	* ** * * * * * * *
<i>C.r.</i>	CTTTCAGATATGAAAAACAATCCGTTTTTAGGTTTTCGCTGGAGCAATTGGGTACTCAATGA
<i>E.c.</i>	CATTTAAATATGAAAAACAACCGTTTTTAGGTTTTGAGGAGCTATTGGTTACTCAATGG
<i>A.m.</i>	CTTTTGCCTTCTCTAAAAAATTAAATCACGTCCTTTCGACGGCGCTGTGGGATATTCTCTGG
	* ** * * * * * * * * *

## FIG. 1B

<i>C.r.</i>	ATGGACCAAGAAATAGAGTTCGAAGTATCCTATGAAACTTTTGATGTAAAAACCTAGGTG	
<i>E.c.</i>	ATGGTCCAAGAAATAGAGCTTGAAGTATCTTATGAAACATTTTGATGTAAAAAATCAAGGTA	
<i>A.m.</i>	GAGGAGCCAGAGTGGAAATTGGAAGCGAGCTACAGAAGTTTGCTACTTTGGCGGACGGGC	** * * * *
		** * * * *
<i>C.r.</i>	GCAACTATAAAAAACAACGCACACATGTACTGTGCTTTAGATACAGCAGCACAATAAGCA	
<i>E.c.</i>	ACAATTATAAGAAATGAAGCACATAGATATTGTGCTCTATCCCATAACTCAGCAGCAGACA	
<i>A.m.</i>	AGTACGCCAAAAAGTG-----GTGGGGAATCTCTGGCAGCTATTACCCGCG	* * * * *
		* * * * *
<i>C.r.</i>	CTAATGGCGCAGGATTAACTACATCTGTATTGGTAAAAAAACGAAATTTAACAAATATAT	
<i>E.c.</i>	TGAGTAGTGCAAG---TAATAATTTGTCTTTCTAAAAAAATGAAGGATTACTTGACATAT	
<i>A.m.</i>	ACGCTAACATTACTGAGACCAATTACTTCGTAGTCAAAATTTGATGAAATCACAAACACCT	* * * * *
		* * * * *
<i>C.r.</i>	CATTAAATGTTAAATGCGTGTATGATATCATGCTTGATGGAATACCAGTTTCTCCATATG	
<i>E.c.</i>	CATTATGCTGAACGCATGCTATGACGCTAGTAGGCGAAGCATACCTTTTCTCCTTATA	
<i>A.m.</i>	CAGTCATGTTAAATGGCTGCTATGACGCTGCTGCACACAGATTTACCTGTGTCCCGTATG	** * * * *
		** * * * *
<i>C.r.</i>	TATGTGCAGGTATTGGCACTGACTTAGTGTGTCAGTAATTAATGCTACAAATCCTAAATTAT	
<i>E.c.</i>	TATGGCGCAGGTATCGGTACTGATTAGTATCCATGTTGAAGCTACAAATCCTAAATTT	
<i>A.m.</i>	TATGTGCCGGGATAGGCGCAAGCTTTGTTGACATCTCTAAGCAAGTAACCAACAAGCTGG	*** * * * *
		*** * * * *
<i>C.r.</i>	CTTATCAAGGAAAGCTAGGCATAAGTTACTCAATCAATCTGAAGCTTCTATCTTTATCG	
<i>E.c.</i>	CTTACCAAGGAAAGTTAGGTTTAAAGCTACTCTATAAGCCAGAGCTTCTGTGTTTATTG	
<i>A.m.</i>	CCTACAGGGGCAAGGTTGGGATTAGCTACCAAGTTACTCCGGAATATCCTTGGTGGCAG	* * * * *
		* * * * *

## FIG. 1C

<i>C.r.</i>	GTGGACATTTCATAGAGTTATAGGTAATGAATTTAAAGATATTGCTACCTTAAAAATAT	
<i>E.c.</i>	GTGGGCACTTTCATAAGGTAATAGGGAACGAATTTAGAGATATTCTACTATAATACCTA	
<i>A.m.</i>	GTGGGTTCTACCAACGGGCTATTTTGATGAGTCTTACAAGGACATTCCTCGCACACAACAGTG	
	*** * ** * * * * * * * * * *	
<i>C.r.</i>	TTACTTCAAAAACAGGAATATCTAATCCTGGCTTTGCATCAGCAACACTTGATGTTGTC	
<i>E.c.</i>	CTGGATCAACACTTGCAGGAAAAGGAACTACCTGCAATAGTAATACTGGATGATGCC	
<i>A.m.</i>	TAAAGTTCTCTGGAGAGCAAAA-----GCCTCAGTCAAAAGCGCATATTGCTG	
	* * * * *	** * * *
<i>C.r.</i>	ACTTTGGTATAGAAAATTGGAGGAAGGTTTGTATTTTAA----	
<i>E.c.</i>	ACTTTGGAATAGAAAATGGGAGGAAGGTTTAA-----	
<i>A.m.</i>	ACTACGGCTTTTAACTTGGAGCAAGATTCTCTGTTTCAGCTAA	
	*** ** * * * **** ** *	

1 ggaatgaaattcaggggacattttctactctttaaagcgtttgctacaccatcattctgcagctta  
N E F R D I S T L K A F A T P S S A A T  
61 ctccagacttagcaacagctaacactgagtgctgtctactctggagtgagaactctggaggtaa  
P D L A T V T L S V C H F G V E L G G R  
121 gattttaaactcttaactttttattattgccacatgtttaaaaataatctaaactctgttttcatt  
F N F \*  
181 attgcttacaataataataaaatagttggcaaaaaaaatgttagcaataaagagggggggggggga  
241 ctaattactatcttgcctatctcccttactaccacttacacataaataatcttgacaaataacaa  
301 cagcttctggagaaataaacaacattttaaaatctttctacaaaaaccattttatcaatcttgt  
-35  
361 actaaaaaactagctttatattctgtttttacattgttaggtttactactgttaattttgtttt  
-10  
421 cactattttcaggtgttaatatgaactgcgaaaaaattttttacaacaactgcatttaacatt  
RBS M N C E K F P I T T A L T L  
481 ctaatgtctcttcttaccctggaaatctcaactcttgatctcagcaccagtagcaacattagt  
L M S F L P G I S L S D P V Q D D N I S  
541 ggaatgtttcacatcagctgaggaagtatatgccaaagcgtcttcgcatctttggagattttttct  
G N F Y I S G K Y M P S A S H F G V F S  
601 gccaaagggaagaaagaaatacaaacagcttggaagtatttggaatagagcaagattggggataga  
A K E E R N T T V G V F G I E Q D W D R  
661 tgtgttaattctagaaccacttttaagcagatatactaccgttcccaattttatcccttaag  
C V I S R T T L S D I F T V P N Y S F K  
721 catcgaataatactattttcaggattttcaggaggtcttggtctactcaatcgagtgggccca  
Y E N N L F S G P A G A I G Y S M D G P  
781 agaactagagcttggaagtactcttatgaagcattcgatgttaaaaactcaaggtaaacattat  
R I E L E V S Y E A F D V K N Q G N N Y  
841 aagaacgaagcagatagatattatgtcttctccacttctctcggcacagagacagagata  
K N E A H R Y Y A L S H L G T E T Q I  
901 gatggtgcaggcagctgcgtctgtcttttaataaatgaaggactacttgataaaatcattt  
D G A G C S A S V P L I N E G L L D K S P  
961 atgtgaacgcatcttatgtatgaataagtgaaagcagatccattctctctcttatatattgt  
M L N A C Y D V I S E G I P P S P Y I C  
1021 gcaggtatttggtatttgatttagtattccatcttgaagcttataaaactcaaaatcttttat  
A G I G I D L V S M P E A I N P K I S Y  
1081 caaggaaaatttaggtcttaagttaccctataagcccagaagctctctgtttttattggtgga  
Q G K L G L S Y P I S P E A S V F I G  
1141 catttctataaggttgataggaagaagaatttagagatacttactatgatcacatagtgaa  
H F H K V I G N E F R D I P T M I P S E  
1201 tcagcgtctgcaggaaaggaactaccctgcaatagtaaacactggagcgtgtcttacttt  
S A L A G K G N Y P A I V T L D V F Y P  
1261 ggcatagaactctggaggaaggttttaacttccaaactctgattattggcacaataaataaaa  
G I E L G G R F N F Q Q L \*  
1321 atagtggcaaaaagaatgtagcaataaagagggggggggggggaactaaattattatttgc  
1381 atatcccttactaccacttacaccaataatctgacaaaatacaacagttcaaacaaaggt  
1441 aaacaattctttaaattgtcttatgagaaccattgttttttatattaaaaacttagetta  
-35  
1501 taactctgtctttacattgcagttctactattgttaatttttttactatttttaggtgta  
-10 RBS  
1561 atatgaattgcaaaaaaattttttataaacaactgcatttagtatcactaatgtctctttctac  
M N C K K P P I T T A L V S L M S F L P  
1621 ctggaatatctattcttgatccagctgcaaggtgacaataatagtggtaatttttattgttta  
G I S F S D P V Q G D N I S G N F Y V S  
1681 gtggcaagtatatgccaaagctcttcgcatctttggcattgtttctgccaaagaagaaaaaa  
G K Y M P S A S H F G M P S A K E E K N  
1741 atcctactgttgcatgtatggcttaaaacaagattgggaagggatagctcactcaagtc  
P T V A L Y G L K Q D W E G I S S S S H  
1801 acaatgataaatctttcaataacaagggcttatctcatataaataagaaataaaccatttt  
N D N H P N N K G Y S F K Y E N N P F L  
1861 taggggtctgcaggagctatttggtttatctaatgggtgggtccagagtagagttggaagtgt  
G F A G A I G Y S M G G P R V E F V S  
1921 cctatgaaacatttgacgtttaaaatcagggtataaactataaaaatgatgctccacagat  
Y E T P D V K N Q G N N Y K N D A H R Y  
1981 actgtgtcttaggtcaacaagacacagcgggaatcccaaaactagtaaatcgttactgt  
C A L G Q Q D N S G I P K T S K Y V T L  
2041 taaaaagcgaaggtatttgacatatcttatgtctaaatgcattgctatgatataataa  
K S E G L L D I S F M L N A C Y D I I N  
2101 acgagagcataccttgtctcttcatatgtgctagggtgttggtactgtatttaataacca  
E S I P L S P Y I C A G V G T D L I S M  
2161 tcttttgaacttacaatactcaaaatctttaccaaggggaagttaggtctaatgtactctta  
F E A T N P K I S Y Q G K L G L S Y S I  
2221 taaaaccgaagctctgtatttttggtggacattttcataaggttgataggaaacgaat  
N P E A S V F I G G H F H K V I G N E P  
2281 ttaggagcatcttacttgaagcatttggttacgtcatcagctctccagatctagcaa  
R D I P T L K K A P V T S S A T P D L A I

FIG. 2A

2341 tagtaacactaagtgatgtcatttttgaatagaacttggaggaagggttaacttttaatt  
 V T L S V C H F G I E L G G R F N F \*  
 2401 tttgttactggccacatgttaaaaaataacttaaaacttgtttttcattattgctacagtaaat  
 2461 aaaaaatagtggtgaaagaatgtagcaataagaaggggggggggactaaattgctatttt  
 2521 accatataccctttattataccacttacactaaataacttgcacaaatacaacagcttttggga  
 2581 aaaacaaacatactttaaacttctctttacaaaaaccatttttttctttgtactaaaaacca  
 -35  
 2641 gctttataacttgtttttacattgtagttctactattgttaattttatttttactatttttag  
 -10  
 2701 gtgcataatcgaattgcacaaaaatttttatacaactacattagttatcgttaattgttttt  
 RBS M N C K K F F I T T T L V S L M S P  
 2761 cttacctgggaatatcattttttgtatgcagtacagaacgacaaatgttgggtggaattttta  
 L P G I S F S D A V Q N D N V G G N F Y  
 2821 tatcagtggggaaatcgtaccaagtggtttcacattttggcggtattctctgctaaacagga  
 I S G K Y V P S V S H F G V P S A K Q E  
 2881 aagaaatatacaaatcggagctatttggattaaagcaagattgggatggcagcacaatate  
 R N T T I G V F G L K Q D W D G S T I S  
 2941 caaaaatttctccagaaaatacatttaacgttcccaaaatttttttttttttttttttttttttt  
 K N S P E N T P N V P N Y S F K Y E N N  
 3001 tccattttttaggttttgcaggagctgttgggttttttaataatgaaatgggtccaaagatagagt  
 P F L G P A G A V G Y L M N G P R I E L  
 3061 agaaatgttcttgaacacatttgcagtggaacacagggttaataactataagaacgatgc  
 E M S Y E T F D V K N Q G N N Y K N D A  
 3121 tcacaaatatttatgcttttaaccataaacagtggtggggaagcttaagcaatgcagggtgataa  
 H K Y Y A L T H N S G G K L S N A G D K  
 3181 gtttgttttttttaaaaaatgaaggactacttgatatactactttatgttgaatgcaggtta  
 F V F L K N E G L L D I S L M L N A C Y  
 3241 tgatgtaataagtgaaaggaataccttttctctctttacatatgtgcagggtgttgggtactga  
 D V I S E G I P P S P Y I C A G V G T D  
 3301 ttttaatatccatgttttgaagctatacaagcttaaaaatttttttttttttttttttttttttt  
 L I S M F E A I N P K I S Y Q G K L G L  
 3361 gagttacttccataagcccagaagcttctgtttttgttgggtggacattttcataaggtgat  
 S Y S I S P E A S V F V G G H F H K V I  
 3421 agggaaatgaattcagagatatctctgttatgataccagtagtcccaacttctcacaggtaa  
 G N E P R D I P A M I P S T S T L T G N  
 3481 tcactttacttatagtaacacataagtgatgccactttggagtggaacttggaggaaggtt  
 H F T I V T L S V C H F G V E L G G R F  
 3541 taacttttttaatttttatttggccacatgttaaaaaataatctaaacttgtttttttattattg  
 N P: \*  
 3601 ctgcaggtaaatataaataagtggtgcaaaagaatgtagcaataacaggggggggggggacttag  
 3661 tttataagtgctgttttttttccactttacacatgatactataacttaaccagtttttttgc  
 3721 tactacttacctgacgtaataatattaaatttttcttacaagaattacagatattttatcac  
 -35  
 3781 aaaaaattttatttttgaacttgcctttttatatgacacttctactattgttaattttattttgct  
 -10  
 3841 actatttaggttttatatgaattacaaaaaagtttttcatacaagtgcatgtatattatta  
 RBS M N Y K K V F I T S A L I S L  
 3901 atactttctctacctggagtatcatttttccgacccagcagggtagtggtatttaacggtaatt  
 I S S L P G V S P S D P A G S G I N G N  
 3961 ttctacatcagtggaataatcatgccaagtgcttgcattttggaggtattctctgctaaag  
 F Y I S G K Y M P S A S H P G V F S A K  
 4021 gaagaaagaaatacaacagttggagtggttggactgaagcaaaattgggacggaagcgca  
 E E R N T T V G V F G L K Q N W D G S A  
 4081 atatcccaacttctcccaaacgagtgatttcaactgtctcaaaatttttcaatttttaataatcaa  
 I S N S S P N D V P T V S N Y S F K Y E  
 4141 aacaacccgttttttaggttttgcaggagctatttgggttactcaatggatgggtcccaagaata  
 N N P F L G F A G A I G Y S M D G P R I  
 4201 gagcttgaagtattcttatgaacattttgatgtaaaaaatcaaggtaacaattataagaat  
 E L E V S Y E T F D V K N Q G N N Y K N  
 4261 gaagcacatagatactgtgctctatccataactcagcagcagacatgagtagtgcaagt  
 E A H R Y C A L S H N S A A D M S S A S  
 4321 aataattttgtcttttttaaaaaatgaaggattacttgacatatcatttatgctgaacgca  
 N N P V F L K N E G L L D I S F M L N A  
 4381 tgcctatgacgtagtaggcgaaggcataccttttttctcttatatatgcgcagggtatcggt  
 C Y D V V G E G I P P S P Y I C A G I G  
 4441 actgatttagtatccatgtttgaagctacaaatcttcaaaatttttttaccaggaaggtta  
 T D L V S M F E A T N P K I S Y Q G K L  
 4501 ggtttaagctactctataagcccagaagcttctgtgtttatttgggtgggcactttcataag  
 G L S Y S I S P E A S V F I G G H F H K  
 4561 gtaataggggaacgaatttagagataattctactataatactactggatcaacactttgca  
 V I G N E P R D I P T I I P T G S T L A  
 4621 ggaaggggaactaccctgcaatagtaataactggatgtatgccactttgggaatagaaatg  
 G K G N Y P A I V I L D V C H F G I E M  
 4681 gga  
 G

FIG. 2B



```

1  tggtgtaaatatgaaatataaaaaaacttttacagtaactgcattagttattattaacttc
   RBS      M K Y K K T F T V T A L V L L T S
61  ctttacacattttatacctttttatagtcacgacgtgcccagtaacaattcacacacttta
   P T H P I P F Y S P A R A S T I H N F Y
121 cattagtggaataatataatgccaacagcgtcacattttggaattttttcagctaaagaaga
   I S G K Y M P T A S H F G I F S A K E E
181 acaaaagttttactaagggtattagttgggttagatcaacgatttatcacataattattataa
   Q S F T K V L V G L D Q R L S H N I I N
241 caataatgatacagcaaaagagttttaagggttcaaaattatttcatttaaatacaaaaataa
   N N D T A K S L K V Q N Y S F K Y K N N
301 cccattttttaggattttgcaggaggtattgggttatccaataggcaattcaagaatagaact
   P F L G F A G A I G Y S I G N S R I E L
361 agaagttatcacatgaaatatttgatactaaaaacccaggaaacaattatttaaatgacttc
   E V S H E I F D T K N P G N N Y L N D S
421 tcacaaatattgcgcttttatctcatggaagtcacatatgcagtgatggaaatagcggaga
   H K Y C A L S H G S H I C S D G N S G D
481 ttggtacactgcaaaaactgataagtttgtaacttttgaaaaatgaagggttactttgacgt
   W Y T A K T D K P V L L K N E G L L D V
541 ctcattttatgttaaacgcattgttatgacataacaactgaaaaaatgcctttttcactta
   S F M L N A C Y D I T T E K M P F S P Y
601 tatatgtgcagggtattgggtactgtctcatatctatgtttgagacaacacaaaaacaaat
   I C A G I G T D L I S M F E T T Q N K I
661 atcttatcaaggaaagtttaggttttaactatactataaaactcaagagttttctgtttttgc
   S Y Q G K L G L N Y T I N S R V S V F A
721 aggtgggcacttttcataaggtaataggtaatgaatttaagggtattcttactctattacc
   G H F H K V I G N E F K G I P T L L P
781 tgatggatcaaacattaaagtacaacagtcctgcaacagtaacattagatgtgtgccattt
   D G S N I K V Q Q S A T V T L D V C H F
841 cgggttagagattggaagtagatttttttttaatactttctattgtacatgttaaaaata
   G L E I G S R F F F F
901 gtactagtttgccttctggttttataaacgcaagagagaatagtttagtaataaattaga
961 aagtttaaatattagaaaagtcatatgtttttcattgtcattgatactcaactaaaagttag
1021 tataaatgttacttattataaattttacgtagratataaattttcccttacaaaagccac
1081 tagtattttatactaaaagctatacttttggtttgtattttaattttgtatttttactactgt
   -35      -10
1141 taattttacttttactgttttttggtgtaaatatgaattgtaaaaaagtttttcacataagt
   RBS      M N C K K V F T I S
1201 gcattgataatccatatacttctacctaagtctctactcttaaccagtatatggt
   A L I S S I Y P L P N V S Y S N P V Y G
1261 aacagtatgtatggttaatttttatacatcaggaaagtacatgccaagtgttcttctttt
   N S M Y G N F Y I S G K Y M P S V P H F
1321 ggaattttttcagctgaagaagagaaaaaaagacaactgtagtatatggcttaaaagaa
   G I F S A E E E K K K T T V V Y G L K E
1381 aactgggcaggagatgcaatattctagtcaaagtcagatgataattttaccatttcgaaat
   N W A G D A I S S Q S P D D N F T I R N
1441 tactcattcaagtatgcaagcaacaagtttttaggggtttgcagtagctattgggttactcg
   Y S F K Y A S N K F L G F A V A I G Y S
1501 ataggcagtcacaagaatagaagttgagatgtcttatgaagcatttgatgtaaaaaatcaa
   I G S P R I E V E M S Y E A P D V K N Q
1561 ggtaacaatt
   G N N

```

FIG. 2C

1 acatgtatacattatagtaacaaatgttacggtattttattcataagttaagtaaaatct  
61 ataccattctctttcactttatcagaagactttttatttatcacaaactcatgacgtatag  
121 tgtcacaaataaacacactgcaactgcaatcactacgtaaaaactttaactcttctttttc  
181 acaactaaaataactaataaaagtaatatagttataaaaaatcttaagtaacTTGACATAat  
-35  
241 attactctgataTAGCATatgtctagtatctctataactaaacgtttatataattGGAGca  
-10  
301 tattaATGAAAGCTATCAAATTCATACTTAATGTCTGCTTACTATTTGCAGCAATATTTT  
M K A I K F I L N V C L L F A → A I F L  
361 TAGGGTATTCCTATATTACAAAACAAGGCATATTTCAAACAAAACATCATGATACACCTA  
G Y S Y I T K Q G I F Q T K H H D T P N  
421 ATACTACTATAACCAATGAAGACGGTATTCAATCTAGCTTTAGCTTAATCAATCAAGACG  
T T I P N E D G I Q S S F S L I N Q D G  
481 GTAAACAGTAACCAAGCAAGATTTCTAGGGAAACACATGTTAGTTTTGTTGGATTCT  
K T V T S Q D F L G K H M L V L F G F S  
541 CTGCATGTAAAGCATTGCCCCTGCAGAATTGGGATTAGTATCTGAAGCACTTGCACAAC  
A C K S I C P A E L G L V S E A L A Q L  
601 TTGGTAATAATGCAGACAAATTACAAGTAATTTTTATTACAATTGATCCAAAAATGATA  
G N N A D K L Q V I F I T I D P K N D T  
661 CTGTAGAAAAATTAAAAGAATTCATGAACATTTTGATTCAAGAATTCAAATGTTAACAG  
V E K L K E F H E H F D S R I Q M L T G  
721 GAAATACTGAAGACATTAATCAAATAATTAAAAATTATAAAATATATGTTGGACAAGCAG  
N T E D I N Q I I K N Y K I Y V G Q A D  
781 ATAAAGATCATCAAATTAACCATCTGCAATAATGTACCTTATTGACAAAAAGGATCAT  
K D H Q I N H S A I M Y L I D K K G S Y  
841 ATCTTTCACACTTCATTCCAGATTTAAAATCACAAGAAAATCAAGTAGATAAGTTACTAT  
L S H F I P D L K S Q E N Q V D K L L S  
901 CTTTAGTTAAGCAGTATCTGTAAtttaataattaattAAAGagaatagtacacaCTTTtt  
L V K Q Y L \*  
961 ataaattcatggaatacgttggtatgagtaggttttttttagtatttttagtgctaataac  
1021 attggcat

FIG. 3A

1 ggaaatctcatgtaaactgaaatactatattcttttttaataaccaatacaattgaata  
61 caaaaaaactttttacaacttattatgtttatctttaaacccttattttaagattccttatg  
121 tcacaaaataacaaaaaactattttacaaaatacaccacaatttcacaaataaaaaaaaa  
181 ctatacacttttattatactacagtagatataccataaaagattttaagtaacTTGACata  
241 atattaccttggtatTAGCATatgattcagttattttatattaaaatttattatgtattGGA  
-10  
301 GcataaaATGAAAGTTATCAAATTTATACTTAATATCTGTTTATTATTTGCAGCAATTTT  
M K V I K F I L N I C L L F A →A I F  
361 TCTAGGATATTCCTACGTAACAAACAAGGCATTTTCAAGTAAGAGATCATAACACTCC  
L G Y S Y V T K Q G I F Q V R D H N T P  
421 CAATACAAATATATCAAATAAAGCCAGCATTACTACTAGTTTTTCGTTAGTAAATCAAGA  
N T N I S N K A S I T T S F S L V N Q D  
481 TGGAAATACAGTAAATAGTCAAGATTTTTTGGGAAAATACATGCTAGTTTTATTGGATT  
G N T V N S Q D F L G K Y M L V L F G F  
541 TTCTTCATGTAAAAGCATCTGCCCTGCTGAATTAGGAATAGCATCTGAAGTTCTCTCACA  
S S C K S I C P A E L G I A S E V L S Q  
601 GCTTGGTAATGACACAGACAAGTTACAAGTAATTTTCATTACAATTGATCCAACAAATGA  
L G N D T D K L Q V I F I T I D P T N D  
661 TACTGTACAAAAATTAAAAACATTTTCATGAACATTTTGATCCTAGAATTCAAATGCTAAC  
T V Q K L K T F H E H F D P R I Q M L T  
721 AGGCAGTGCAGAAGATATTGAAAAAATAATAAAAAATTACAAATATATGTTGGACAAGC  
G S A E D I E K I I K N Y K I Y V G Q A  
781 AGATAAAGATAATCAAATTGATCACTCTGCCATAATGTACATTATCGATAAAAAAGGAGA  
D K D N Q I D H S A I M Y I I D K K G E  
841 ATACATTTACACTTTTCTCCAGATTTAAATCAACAGAAAATCAAGTAGATAAGTTACT  
Y I S H F S P D L K S T E N Q V D K L L  
901 ATCTATAATAAAACAATATCTCTAAtttaataattaattaAAGAGaatagtagacacaCTCT  
S I I K Q Y L \*  
961 Tatataaattcatggatatatgtgatgggtagatttcttttggtgtttctatcgctaatt  
1021 acatta

FIG. 3B

# INTERNATIONAL SEARCH REPORT

Inte. onal Application No

PCT/US 97/19044

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C07K14/29 C12N15/86 A61K31/70

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MCGUIRE T. C. ET AL.: "Recombinant vaccinia virus expression of anaplasma marginale surface protein MSP-1a: effect of promoters, leader sequences and GPI anchor sequence on antibody response" VACCINE, vol. 12, no. 5, - 1994 pages 465-471, XP002057342	1,2,6,7, 10,11
Y	see the whole document	3,4,12, 13,16
Y	--- OBERLE S. M. & BARBET A.F.: "Derivation of the complete msp4 gene sequence of anaplasma marginale without cloning" GENE, vol. 136, - 1993 pages 291-294, XP002057343 see whole document; esp. p293, par. d ff --- -/--	3,4,12, 13,16

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

2 March 1998

Date of mailing of the international search report

19. 03. 1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Müller, F

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 97/19044

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 90 12030 A (UNIV WASHINGTON) 18 October 1990 see whole doc, esp. claims 61-63 ---	1,2
X	VAN VLIET A.H.M. ET AL.,: "Molecular cloning, sequence analysis and expression of the gene encoding the immunodominant 32-kilodalton protein of cowdria ruminantium" INFECT. AND IMMUNITY, vol. 62, no. 4, - April 1994 pages 1451-1456, XP002057344 see the whole document -----	1,2,10, 11,19

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 97/ 19044

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 10-18  
because they relate to subject matter not required to be searched by this Authority, namely:  
Remark: Although claims 10-18 are directed to a method of treatment of the human/animal body the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 97/19044

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9012030 A	18-10-90	AU 5521490 A	05-11-90
		EP 0467972 A	29-01-92
		JP 4504422 T	06-08-92
		US 5549898 A	27-08-96
-----			